UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/662,602	09/15/2003	Chang-Ning Huang	M61.12-0514	2369

27366   7590   09/04/2008
WESTMAN CHAMPLIN (MICROSOFT CORPORATION)
SUITE 1400
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402-3244

EXAMINER
SAINT CYR, LEONARD

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
09/04/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)
Office Action Summary	10/662,602	HUANG ET AL.
	Examiner	Art Unit	
	LEONARD SAINT CYR	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *30 May 2008*.

2a) ☒ This action is **FINAL**.   2b) ☐ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *27,31, 32, and 34 - 46* is/are pending in the application.

     4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *27,31, 32, and 34 - 46* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☒ The drawing(s) filed on *09/15/03* is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

     a) ☐ All   b) ☐ Some *   c) ☐ None of:

         1. ☐ Certified copies of the priority documents have been received.

         2. ☐ Certified copies of the priority documents have been received in Application No. _____.

         3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

     * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Response to Arguments

1.   Applicant's arguments filed 05/30/08 have been fully considered but they are not

persuasive.

Applicant argues that Wu does not teach or suggest generating a quantitative

value that represents the precision with word type indications were applied in output

generated based on a language model being evaluated (Amendment, pages 6 - 8).

The examiner disagrees, Wu discloses "this scoring tool measures word recall

word precision, the F-measure…Given a reference (the gold standard) and a hypothesis

(the segmentation hypothesized by the word segmenter)...word precision is the

percentage of words in the hypothesis that are also in the reference. Most common

types of named entities and factoids whose segmentation may vary across different

standards" (page 19, lines 1 – 7; pages 12, and 13). Measuring a word precision that is

the percentage of words in the segmentation hypothesized by the word segmenter and

the reference implies generating a quantitative value that represents the precision with

word type indications were applied in output generated based on a language model

being evaluated, since segmentation of most common types of named entities and

factoids is varied across different standards.

Claim Rejections - 35 USC § 102

2.   The text of those sections of Title 35, U.S. Code not included in this action can

be found in a prior Office action.

3.   Claims 27 - 33 are rejected under 35 U.S.C. 102(e) as being anticipated by Wu

(Customizable Segmentation of Morphologically Derived Words in Chinese; February

2003).

As per claim 27, Wu teaches a method for evaluating a word segmentation

language model, comprising:

building the word segmentation language model based on an annotated corpus;

applying the language model to a test corpus of unsegmented text different from the

annotated corpus to provide an output indicative of words in the test corpus, and a word

type indication for each word ("most common types of named entities and factoids

whose segmentation may vary across different standards"), the word type indication

being one of a plurality of word type indications ("a language model is used to select the

best sequence of characters"; page 2, paragraph 2, lines 9 – 11; page 9, paragraph 4,

lines 6, and 7; page 12, sections 2.5);

comparing the word type indication for each word in the output of the language

model with predefined word type indications of words of the test corpus; and generating

a quantitative value that represents a level of precision with which word type indications

were applied in the output indicative of words in the test corpus, wherein generating

comprising generating based on a comparison of the word type indication for words in

the output to the predefined word type indications ("this scoring tool measures word

precision, the F-measure…Given a reference (the gold standard) and a hypothesis (the

segmentation hypothesized by the word segmenter)…word precision is the percentage

of words in the hypothesis that are also in the reference"; page 19, lines 1 – 7; pages

12, and 13).

As per claim 31, Wu teaches a method of evaluating word segmentation models,

comprising:

using a first word segmentation model to segment a corpus of text into words and

tags to the words indicative of one of the plurality of word types ("most common types of

named entities and factoids whose segmentation may vary across different standards"),

the words and tags forming a first output ("first segmented the text using our default

setting"; page 19, paragraph 2, lines 1, and 2; page 12, sections 2.5);

using a second word segmentation model to segment a corpus of text into words

and tags to the words indicative of one of the plurality of word types ("most common

types of named entities and factoids whose segmentation may vary across different

standards"), the words and tags forming a second output ("adjusted segmentation were

evaluated"; page 19, paragraph 2, lines 1 - 4; page 12, sections 2.5);

comparing the first output to a predefined indication of words and tags of the

words indicative of one of the plurality of words types from the corpus of text to provide

a first set of values for each of the plurality of word types indicative of how the first word

segmentation model recognizes each of the plurality of word types; comparing the

second output to a predefined indication of words and tags of the words indicative of

one of the plurality of words types from the corpus of text to provide a second set of

values for each of the plurality of word types indicative of how the second word

segmentation model recognizes each of the plurality of word types ("scores when the

default setting was used…after a quick adjustment…the scores became"; page 19,

paragraph 3);

comparing the first set of values and the second set of values to determine

effectiveness of the first word segmentation model and the second word segmentation

with respect to each of the plurality of word types ("the scores improved dramatically

across the board in both the CHTB and BU data after the parameter values were

adjusted to the relevant standards"; page 20, paragraph 3).

As per claim 32, Wu further discloses that the first set of values is based on

matches between the first output and the predefined indication and wherein the second

set of values is based on matches between the second output and the predefined

indication (page 19; paragraphs 1, and 2).

As per claim 34, Wu further discloses generating the quantitative value

comprises generating a quantitative value based on a comparison of word type

indications of words in the output that match predefined word type indications assigned

to the same words in the test corpus ("word precision is the percentage of words in the

hypothesis that are also in the reference"; page 19, lines 1 – 7).

As per claim 35, Wu further discloses generating the quantitative value comprises generating a quantitative value that is indicative of how frequently a word type indication in the output matches a corresponding predefined word type indication in the test corpus ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs"; page 19, lines 1 – 7; page 24, section 3.5).

As per claim 36, Wu further discloses generating the quantitative value comprises generating a quantitative value that is indicative of how frequently a word type indication, assigned to a word in the output, matches a predefined word type indication assigned to a same word in the test corpus ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs"; page 19, lines 1 – 7; page 24, section 3.5).

As per claim 37, Wu further discloses generating a quantitative value further comprises generating a quantitative value that represents a level of precision with which person name word type indications were applied in the output ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs...personal names"; page 19, lines 1 – 7; page 12, section 2.5.1).

As per claim 38, Wu further discloses generating a quantitative value further comprises generating a quantitative value that represents a level of precision with which location name type indications were applied in the output ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…place names"; page 19, lines 1 – 7; page 12, section 2.5.2).

As per claim 39, Wu further discloses generating a quantitative value further comprises generating a quantitative value that represents a level of precision with which organization name word type indications were applied in the output ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…organization names"; page 19, lines 1 – 7; page 12, section 2.5.2).

As per claim 40, Wu further discloses generating a quantitative value further comprises generating a quantitative value that represents a level of precision with which overlapping ambiguous string word type indications were applied in the output ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…segmentation ambiguities are resolved in the parsing process""; page 19, lines 1 – 7; page 5, lines 1 - 3).

As per claim 41, Wu further discloses generating a quantitative value further comprises generating a quantitative value that represents a level of precision with which covering ambiguous string word type indications were applied in the output ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…segmentation ambiguities are resolved in the parsing process""; page 19, lines 1 – 7; page 5, lines 1 - 3).

As per claim 42, Wu further discloses comparing to provide a first set of values for each of the plurality of word types comprises comparing to provide a first set of values for a person name word type ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…personal names"; page 19, lines 1 – 7; page 12, section 2.5.1).

As per claim 43, Wu further discloses comparing to provide a first set of values for each of the plurality of word types comprises comparing to provide a first set of values for a location name word type ("word precision is the percentage of words in the hypothesis that are also in the reference…the segmentation of certain MDWs can depend on the frequency of those MDWs…place names"; page 19, lines 1 – 7; page 12, section 2.5.2).

As per claim 44, Wu further discloses comparing to provide a first set of values for each of the plurality of word types comprises comparing to provide a first set of values for an organization name word type ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs...organization names"; page 19, lines 1 – 7; page 12, section 2.5.2).

As per claim 45, Wu further discloses comparing to provide a first set of values for each of the plurality of word types comprises comparing to provide a first set of values for a overlapping ambiguous string word type ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs...segmentation ambiguities are resolved in the parsing process""; page 19, lines 1 – 7; page 5, lines 1 - 3).

As per claim 46, Wu further discloses comparing to provide a first set of values for each of the plurality of word types comprises comparing to provide a first set of values for a covering ambiguous string word type ("word precision is the percentage of words in the hypothesis that are also in the reference...the segmentation of certain MDWs can depend on the frequency of those MDWs...segmentation ambiguities are resolved in the parsing process""; page 19, lines 1 – 7; page 5, lines 1 - 3).

Conclusion

4.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

5.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to LEONARD SAINT CYR whose telephone number is

(571) 272-4247. The examiner can normally be reached on Mon- Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone

number for the organization where this application or proceeding is assigned is (571)-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

LS
08/22/08

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626